

RECEIVED

AUG 21 2001

TECH CENTER 1600/2900

#8  
7/21/01

1636

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/679,705

DATE: 07/27/2001

TIME: 13:05:30

Input Set : A:\Hmv05201.app

Output Set: N:\CRF3\07272001\I679705.raw

ENTERED

3 <110> APPLICANT: BECKWITH, JONATHAN  
 4 ASLUND, FREDRIK  
 5 BESSETTE, PAUL H.  
 6 GEORGIU, GEORGE  
 7 RITZ, DANIEL  
 8 LIM, JACKIE EUN-AH  
 10 <120> TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR PRODUCTION OF DISULFIDE  
 11 BOND CONTAINING PROTEINS IN HOST CELLS  
 13 <130> FILE REFERENCE: Hmv-052.01  
 15 <140> CURRENT APPLICATION NUMBER: 09/679,705  
 16 <141> CURRENT FILING DATE: 2000-10-05  
 18 <150> PRIOR APPLICATION NUMBER: 60/157,770  
 19 <151> PRIOR FILING DATE: 1999-10-05  
 21 <150> PRIOR APPLICATION NUMBER: 60/163,939  
 22 <151> PRIOR FILING DATE: 1999-11-08  
 24 <150> PRIOR APPLICATION NUMBER: 60/166,044  
 25 <151> PRIOR FILING DATE: 1999-11-17  
 27 <160> NUMBER OF SEQ ID NOS: 24  
 29 <170> SOFTWARE: PatentIn Ver. 2.1  
 31 <210> SEQ ID NO: 1  
 32 <211> LENGTH: 4  
 33 <212> TYPE: PRT  
 34 <213> ORGANISM: Artificial Sequence  
 36 <220> FEATURE:  
 37 <223> OTHER INFORMATION: Description of Artificial Sequence: Illustrative  
 38 motif  
 40 <220> FEATURE:  
 41 <221> NAME/KEY: MOD\_RES  
 42 <222> LOCATION: (2)..(3)  
 43 <223> OTHER INFORMATION: Any amino acid  
 45 <400> SEQUENCE: 1  
 46 Cys Xaa Xaa Cys  
 47 1  
 50 <210> SEQ ID NO: 2  
 51 <211> LENGTH: 4  
 52 <212> TYPE: PRT  
 53 <213> ORGANISM: Artificial Sequence  
 55 <220> FEATURE:  
 56 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
 57 variant motif  
 59 <400> SEQUENCE: 2  
 60 Cys Gly Pro Cys  
 61 1  
 64 <210> SEQ ID NO: 3  
 65 <211> LENGTH: 4  
 66 <212> TYPE: PRT  
 67 <213> ORGANISM: Artificial Sequence

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/679,705

DATE: 07/27/2001

TIME: 13:05:30

Input Set : A:\Hmv05201.app

Output Set: N:\CRF3\07272001\I679705.raw

69 <220> FEATURE:  
70 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
71 variant motif  
73 <400> SEQUENCE: 3  
74 Cys Gly Ser Cys  
75 1  
78 <210> SEQ ID NO: 4  
79 <211> LENGTH: 4  
80 <212> TYPE: PRT  
81 <213> ORGANISM: Artificial Sequence  
83 <220> FEATURE:  
84 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
85 variant motif  
87 <400> SEQUENCE: 4  
88 Cys Pro Tyr Cys  
89 1  
92 <210> SEQ ID NO: 5  
93 <211> LENGTH: 4  
94 <212> TYPE: PRT  
95 <213> ORGANISM: Artificial Sequence  
97 <220> FEATURE:  
98 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
99 variant motif  
101 <400> SEQUENCE: 5  
102 Cys Pro His Cys  
103 1  
106 <210> SEQ ID NO: 6  
107 <211> LENGTH: 4  
108 <212> TYPE: PRT  
109 <213> ORGANISM: Artificial Sequence  
111 <220> FEATURE:  
112 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
113 variant motif  
115 <400> SEQUENCE: 6  
116 Cys Gly His Cys  
117 1  
120 <210> SEQ ID NO: 7  
121 <211> LENGTH: 4  
122 <212> TYPE: PRT  
123 <213> ORGANISM: Artificial Sequence  
125 <220> FEATURE:  
126 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
127 variant motif  
129 <400> SEQUENCE: 7  
130 Cys Gly Pro Ala  
131 1  
134 <210> SEQ ID NO: 8  
135 <211> LENGTH: 48  
136 <212> TYPE: DNA

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/679,705

DATE: 07/27/2001

TIME: 13:05:30

Input Set : A:\Hmv05201.app

Output Set: N:\CRF3\07272001\I679705.raw

```

137 <213> ORGANISM: Escherichia coli
139 <220> FEATURE:
140 <221> NAME/KEY: CDS
141 <222> LOCATION: (1)..(48)
143 <400> SEQUENCE: 8
144 tgg agc gtc ttc ttc ttc tac ccg gct gac ttt act ttc gta tgc ccg      48
145 Trp Ser Val Phe Phe Phe Tyr Pro Ala Asp Phe Thr Phe Val Cys Pro
146   1             5             10             15
149 <210> SEQ ID NO: 9
150 <211> LENGTH: 16
151 <212> TYPE: PRT
152 <213> ORGANISM: Escherichia coli
154 <400> SEQUENCE: 9
155 Trp Ser Val Phe Phe Phe Tyr Pro Ala Asp Phe Thr Phe Val Cys Pro
156   1             5             10             15
160 <210> SEQ ID NO: 10
161 <211> LENGTH: 51
162 <212> TYPE: DNA
163 <213> ORGANISM: Escherichia coli
165 <220> FEATURE:
166 <221> NAME/KEY: CDS
167 <222> LOCATION: (1)..(51)
169 <400> SEQUENCE: 10
170 tgg agc gtc ttc ttc ttc ttc tac ccg gct gac ttt act ttc gta tgc      48
171 Trp Ser Val Phe Phe Phe Phe Tyr Pro Ala Asp Phe Thr Phe Val Cys
172   1             5             10             15
174 ccg                                                                51
175 Pro
178 <210> SEQ ID NO: 11
179 <211> LENGTH: 17
180 <212> TYPE: PRT
181 <213> ORGANISM: Escherichia coli
183 <400> SEQUENCE: 11
184 Trp Ser Val Phe Phe Phe Tyr Pro Ala Asp Phe Thr Phe Val Cys
185   1             5             10             15
187 Pro
190 <210> SEQ ID NO: 12
191 <211> LENGTH: 32
192 <212> TYPE: PRT
193 <213> ORGANISM: Escherichia coli
195 <400> SEQUENCE: 12
196 Arg Trp Ser Val Phe Phe Phe Tyr Pro Ala Asp Phe Thr Phe Val Cys
197   1             5             10             15
199 Pro Thr Glu Leu Gly Asp Val Ala Asp His Tyr Glu Glu Leu Gln Lys
200   20             25             30
203 <210> SEQ ID NO: 13
204 <211> LENGTH: 32
205 <212> TYPE: PRT
206 <213> ORGANISM: Salmonella typhi

```

## RAW SEQUENCE LISTING

DATE: 07/27/2001

PATENT APPLICATION: US/09/679,705

TIME: 13:05:30

Input Set : A:\Hmv05201.app

Output Set: N:\CRF3\07272001\I679705.raw

```

208 <400> SEQUENCE: 13
209 Arg Trp Ser Val Phe Phe Phe Tyr Pro Ala Asp Phe Thr Phe Val Cys
210   1           5           10           15
212 Pro Thr Glu Leu Gly Asp Val Ala Asp His Tyr Glu Glu Leu Gln Lys
213           20           25           30
216 <210> SEQ ID NO: 14
217 <211> LENGTH: 32
218 <212> TYPE: PRT
219 <213> ORGANISM: Pseudomonas putida
221 <400> SEQUENCE: 14
222 Lys Trp Ser Val Val Phe Phe Tyr Pro Ala Asp Phe Thr Phe Val Cys
223   1           5           10           15
225 Pro Thr Glu Leu Gly Asp Leu Ala Asp Asn Tyr Ala Glu Phe Gln Lys
226           20           25           30
229 <210> SEQ ID NO: 15
230 <211> LENGTH: 32
231 <212> TYPE: PRT
232 <213> ORGANISM: Staphylococcus mutans
234 <400> SEQUENCE: 15
235 Lys Trp Ala Val Phe Cys Phe Tyr Pro Ala Asp Phe Ser Phe Val Cys
236   1           5           10           15
238 Pro Thr Glu Leu Gly Asp Leu Gln Glu Gln Tyr Ala Thr Leu Gln Ser
239           20           25           30
242 <210> SEQ ID NO: 16
243 <211> LENGTH: 32
244 <212> TYPE: PRT
245 <213> ORGANISM: Bacillus subtilis
247 <400> SEQUENCE: 16
248 Gln Trp Ser Val Phe Cys Phe Tyr Pro Ala Asp Phe Ser Phe Val Cys
249   1           5           10           15
251 Pro Thr Glu Leu Glu Asp Leu Gln Glu Gln Tyr Ala Ala Leu Lys Glu
252           20           25           30
255 <210> SEQ ID NO: 17
256 <211> LENGTH: 32
257 <212> TYPE: PRT
258 <213> ORGANISM: Staphylococcus aureus
260 <400> SEQUENCE: 17
261 Ser Trp Ser Val Val Cys Phe Tyr Pro Ala Asp Phe Ser Phe Val Cys
262   1           5           10           15
264 Pro Thr Glu Leu Glu Asp Leu Gln Asn Gln Tyr Glu Glu Leu Gln Lys
265           20           25           30
268 <210> SEQ ID NO: 18
269 <211> LENGTH: 32
270 <212> TYPE: PRT
271 <213> ORGANISM: Treponema pallidum
273 <400> SEQUENCE: 18
274 Ser Trp Ala Val Phe Met Phe Tyr Pro Ala Asp Phe Thr Phe Val Cys
275   1           5           10           15
277 Pro Thr Glu Leu Ala Asp Leu Ala Arg Val Tyr Pro Ser Phe Val Glu

```

## RAW SEQUENCE LISTING

DATE: 07/27/2001

PATENT APPLICATION: US/09/679,705

TIME: 13:05:30

Input Set : A:\Hmv05201.app

Output Set: N:\CRF3\07272001\I679705.raw

```

278          20          25          30
281 <210> SEQ ID NO: 19
282 <211> LENGTH: 32
283 <212> TYPE: PRT
284 <213> ORGANISM: Aquifex aeolicus
286 <400> SEQUENCE: 19
287 Lys Trp Val Ile Leu Phe Phe Tyr Pro Ala Asp Tyr Thr Phe Val Cys
288   1          5          10          15
290 Pro Thr Glu Leu Ala Asp Leu Ala Glu Lys Tyr Asp Glu Leu Lys Glu
291          20          25          30
294 <210> SEQ ID NO: 20
295 <211> LENGTH: 32
296 <212> TYPE: PRT
297 <213> ORGANISM: Homo sapiens
299 <400> SEQUENCE: 20
300 Lys Tyr Val Val Leu Phe Phe Tyr Pro Leu Asp Phe Thr Phe Val Cys
301   1          5          10          15
303 Pro Thr Glu Ile Ile Ala Phe Thr Thr Val Lys Arg Thr Ser Ala Lys
304          20          25          30
307 <210> SEQ ID NO: 21
308 <211> LENGTH: 1483
309 <212> TYPE: DNA
310 <213> ORGANISM: Escherichia coli
312 <220> FEATURE:
313 <221> NAME/KEY: CDS
314 <222> LOCATION: (195)..(758)
316 <400> SEQUENCE: 21
317 aagggtagt cagattacac ggtcacctgg aaagggggcc attttacttt ttatgcgctg 60
319 gcggtgcaaa gtccacaaag ttgtcttacg aaggttgtaa ggtaaaactt atcgatttga 120
321 taatggaaac gcattaccgg aatcggcaaa aattggttac cttacatctc atcgaaaaca 180
323 cggaggaagt atag atg tcc ttg att aac acc aaa att aaa cct ttt aaa 230
324          Met Ser Leu Ile Asn Thr Lys Ile Lys Pro Phe Lys
325          1          5          10
327 aac cag gca ttc aaa aac ggc gaa ttc atc gaa atc acc gaa aaa gat 278
328 Asn Gln Ala Phe Lys Asn Gly Glu Phe Ile Glu Ile Thr Glu Lys Asp
329          15          20          25
331 acc gaa ggc cgc tgg agc gtc ttc ttc ttc tac ccg gct gac ttt act 326
332 Thr Glu Gly Arg Trp Ser Val Phe Phe Phe Tyr Pro Ala Asp Phe Thr
333          30          35          40
335 ttc gta tgc ccg acc gaa ctg ggt gac gtt gct gac cac tac gaa gaa 374
336 Phe Val Cys Pro Thr Glu Leu Gly Asp Val Ala Asp His Tyr Glu Glu
337          45          50          55          60
339 ctg cag aaa ctg ggc gta gac gta tac gca gta tct acc gat act cac 422
340 Leu Gln Lys Leu Gly Val Asp Val Tyr Ala Val Ser Thr Asp Thr His
341          65          70          75
343 ttc acc cac aaa gca tgg cac agc agc tct gaa acc atc gct aaa atc 470
344 Phe Thr His Lys Ala Trp His Ser Ser Ser Glu Thr Ile Ala Lys Ile
345          80          85          90
347 aaa tat gcg atg atc ggc gac ccg act ggc gcc ctg acc cgt aac ttc 518

```

**VERIFICATION SUMMARY**

PATENT APPLICATION: US/09/679,705

DATE: 07/27/2001

TIME: 13:05:31

Input Set : A:\Hmv05201.app

Output Set: N:\CRF3\07272001\I679705.raw

L:46 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1